Forecasting Tropical Cyclone Eye Formation and Dissipation in Infrared Imagery

JOHN A. KNAFF

NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado

ROBERT T. DEMARIA

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado

(Manuscript received 21 March 2017, in final form 13 July 2017)

ABSTRACT

The development of an infrared (IR; specifically near 11 μm) eye probability forecast scheme for tropical cyclones is described. The scheme was developed from an eye detection algorithm that used a linear discriminant analysis technique to determine the probability of an eye existing in any given IR image given information about the storm center, motion, and latitude. Logistic regression is used for the model development and predictors were selected from routine information about the current storm (e.g., current intensity), forecast environmental factors (e.g., wind shear, oceanic heat content), and patterns/information (e.g., convective organization, tropical cyclone size) extracted from the current IR image. Forecasts were created for 6-, 12-, 18-, 24-, and 36-h forecast leads. Forecasts were developed using eye existence probabilities from North Atlantic tropical cyclone cases (1996–2014) and a combined North Atlantic and North Pacific (i.e., Northern Hemisphere) sample. The performance of North Atlantic–based forecasts, tested using independent eastern Pacific tropical cyclone cases (1996–2014), shows that the forecasts are skillful versus persistence at 12–36 h, and skillful versus climatology at 6–36 h. Examining the reliability and calibration of those forecasts shows that calibration and reliability of the forecasts is good for 6–18 h, but forecasts become a little overconfident at longer lead times. The forecasts also appear unbiased. The small differences between the Atlantic and Northern Hemisphere formulations are discussed. Finally, and remarkably, there are indications that smaller TCs are more prone to form eye features in all of the TC areas examined.

1. Introduction

The formation of an eye feature in tropical cyclones (TCs) in visible and infrared (IR) imagery is one of the more fascinating events to witness in real time. The transition can occur intermittently, via banding, or rather abruptly, and has distinct forecast implications. Persistent eye formation, which occurs in IR imagery when the TC first reaches ~79 kt (1 kt = 0.514 m s^{-1}) (Vigh et al. 2012), is important to both subjective (Dvorak 1984) and objective (e.g., Olander and Velden 2007) TC intensity estimates. In fact, the formation of an eye typically results in at least a 15-kt increase in intensity, and if the eye becomes increasingly well developed, can require the Dvorak analysts to break constraints to provide a more accurate intensity estimate (Velden et al. 2006). It is important to note that eye formation is often thought to be the beginning of more rapid intensification and that eye formation typically occurs in radar at intensities around 58 kt (28 m s^{-1}) (Vigh et al. 2012 and references within). There is also compelling evidence that the efficiency of tropical cyclones, defined as the ratio of injected heat energy to the kinetic energy change of the symmetric vortex (Schubert and Hack 1982), improves with vortex intensity (inertial stability) and Coriolis parameter with the majority of the increases in efficiency coming from intensification (Nolan et al. 2007). In the same manner, the disintegration of an eye feature is typically related to weakening and to both structural changes, like extratropical transition [i.e., those described by Evans and Hart (2003)], and those caused by poorer kinematic and thermodynamic environments. The disintegration of the eye can also be related to rapid weakening (cf. Wood and Ritchie 2015) if the TC is initially very intense and has a distinct eye.

While it is recognized that operational TC centers do not typically forecast the formation of eye features, it is also documented that there are robust relationships

Corresponding author: John Knaff, john.knaff@noaa.gov

DOI: 10.1175/WAF-D-17-0037.1

between eye formation and rapid intensification (e.g., Malkus 1958; Holliday and Thompson 1979; Weatherford and Gray 1988; Willoughby 1990). Thus, our thesis is that if one was able to anticipate eye behavior, such forecasts may ultimately be very helpful for making short-term (≤36 h) TC intensity forecasts. These forecasts could also be used in combination with the existence of eye features in passive microwave imagery, which often appears earlier than the IR eye feature (Burton et al. 2010), and can be used to forecast rapid intensification (Rozoff et al. 2015). These short-term forecasts are also the time periods that have shown the least improvement over the last several decades (DeMaria et al. 2014) and making improvements in forecasting rapid intensification is one of the highest priorities at U.S. operational centers (OAR 2017). With this premise in mind, the development of such forecasts for the probability of existence of an eye feature in infrared imagery is the topic of this paper. The data and methods are described and results are presented in the next two sections, followed by an example, the development of a Northern Hemisphere model, and a short summary and discussion.

2. Data and methods

The developmental basis for forecasting TC eye formation/dissipation comes from the output of an IR imagery–based eye probability estimate or eye detection algorithm (EDA) developed by DeMaria (2016, hereafter D16), which is now briefly described. Eye probabilities were estimated using linear and quadratic discriminant analysis (LDA and QDA, respectively). The training data for whether a scene had an eye was based on routine Dvorak intensity fixes (following Dvorak 1984), available from the Satellite Analysis Branch (SAB) and the Tropical Forecast and Analysis Branch (TAFB), that are archived in the databases of the Automated Tropical Cyclone Forecast system (ATCF; Sampson and Schrader 2000) and the CIRA tropical cyclone IR (~11 μm) image archive. The latter contains 4 km × 4 km resolution images collected from the global constellation of geostationary satellites and is remapped onto a common Mercator projection that has approximately 30-min temporal resolution. Details of how that eye and no-eye information is extracted from the individual Dvorak fixes are described in the appendix. For this study and D16, TAFB was used as the true eye/no-eye observation.

The IR-based inputs to the EDA are principle components (PCs) of the 320 km × 320 km storm-centered IR imagery. Principle component analysis significantly reduces the dimension of this problem and allows for the use of a moderately sized developmental dataset. Initially, 25 PCs were examined, but their predictive ability leveled off when 10 or more PCs were used. For this reason, a total of 10 PCs with the highest magnitude in the significance vector (i.e., those that contribute the most toward improving the estimate) were used along with four components that come from the best tracks, including the current intensity, zonal storm motion, meridional storm motion, and storm latitude (D16). It is important to note that the best-track information is also routinely available as part of the TC advisories and forecasts. It is also worth mentioning that the largest predictive contributions come from the PCs with eigenvectors that portrayed symmetric/eyelike patterns in the IR imagery, which includes the first 5 of the 10 PCs used, as shown in D16.

Testing and development of the EDA was conducted using 4109 cases from 265 different Atlantic TCs during 1995–2013. All cases are 6 h or greater apart in time, and 991 (24%) were considered to have an eye by TAFB. SAB fixes were not used for development, but eye classifications from SAB and TAFB agreed 95% of the time (D16). To evaluate the quality of the eye detection, the data samples were randomly shuffled and partitioned so that 70% of the data would be used for training and 30% would be used for testing. This process was performed 1000 times (D16).

The resulting algorithms allow for the estimation of the probability of an eye existing in any storm-centered image. Figure 1 shows an example time series of the LDA and QDA versions of the EDA and corresponding 6-hourly IR imagery for east Pacific Hurricane Sandra (2015), which was described as the latest forming east Pacific major hurricane by Stewart (2016). This case was chosen for its short best track and rather typical intensification to major hurricane intensity and decay over cooler waters. Sandra had a pinhole-eye feature (0600 UTC 26 November), reached 130-kt intensity, and was relatively short lived (Stewart 2016). It is clear the LDA and QDA produce different results and disagree primarily at the formation and disintegration times. In the most basic sense this is because the QDA is more sensitive to the image information and less sensitive to the best-track information. In general, the QDA is a bit noisier, fluctuating more rapidly between images, and LDA is more conservative. For our work here, we use the LDA version of the EDA and henceforth EDA will refer to the LDA version of the EDA, which produces the best overall statistics that result in 90% of the cases being correctly predicted, Brier skill scores of 61% (climatology ~ 12%), and Pierce skill scores of 66%; all of which suggests that the EDA can detect IR eye features, can outperform climatology, has skill in discriminating eye

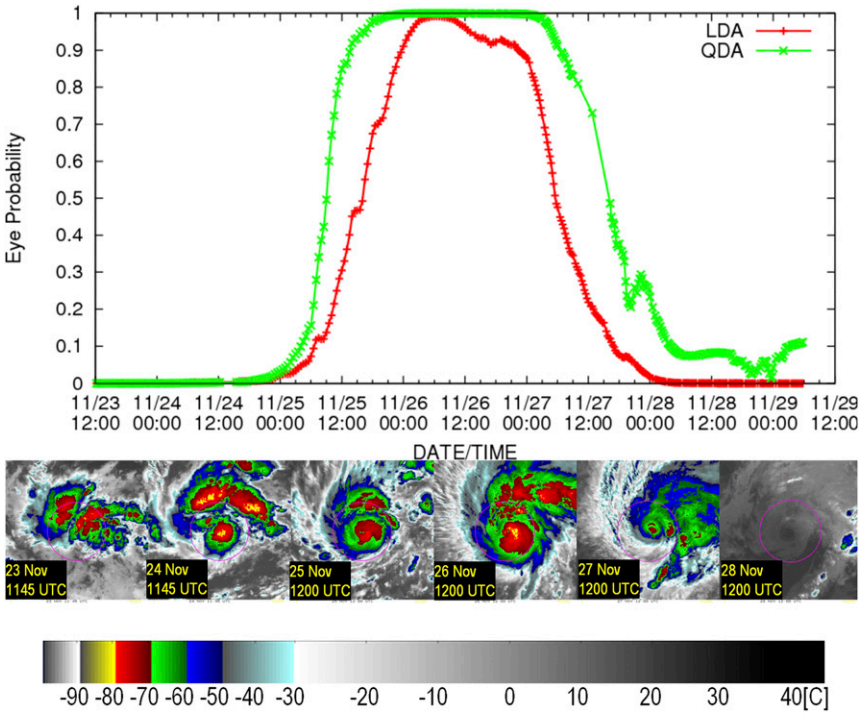

FIG. 1. (top) Comparison of the LDA and QDA versions of the EDA for east Pacific Hurricane Sandra (2015), (bottom) along with daily IR images.

versus no-eye scenes, and is commonly less noisy from one image to another. It is however important to note that the algorithm is rather conservative to rapid changes in the eye and will not necessary detect (probability > 0.5) weak or intermittent eye scenes, just like a conservative human analyst. Full details of both of the EDAs are available in D16.

Having a skillful method to estimate whether a TC image contains an eye allows one to visualize and track eye development and decay. The existence of such a dataset allows us to construct and verify simple statistical forecast models to anticipate eye formation in the future. In addition to the information in the EDA (providing the dependent variables), we would like to incorporate information about the future environment and initial conditions in the IR imagery. Information in the IR imagery has been shown to be related to short-term changes in TC intensity, as demonstrated by DeMaria et al. (2005), Kaplan et al. (2010), and Kaplan et al. (2015), and TC structure (Knaff et al. 2014, 2016; Kossin et al. 2007; Mueller et al. 2006), the latter also being related to intensification rates (e.g., Carrasco et al. 2014). Several environmental conditions are also known to be related to TC intensity changes including vertical wind shear, midlevel moisture, oceanic heat content, and sea surface temperatures, as has been documented by many statistical–dynamical methods for TC intensity forecasting (DeMaria and Kaplan 1999; Knaff et al. 2005; DeMaria et al. 2005; DeMaria 2009) and forecasting rapid intensity change (Kaplan and DeMaria 2003; Kaplan et al. 2010, 2015).

In this work, a similar approach to rapid intensity indices (e.g., Kaplan et al. 2015) is used, that is, employing a combination of current TC conditions, environmental conditions, and information about the current IR structure of the TC to forecast eye probabilities. We however use the method of logistic regression exclusively to provide probabilistic forecasts of IR eye existence at 6, 12, 18, 24, and 36 h.

Logistic regression is a regression model where the dependent variable (DV) is categorical; in this case 1 indicates having an IR eye (p_e) and 0 not having an IR eye (p_n). Logistic regression is also a special case of the generalized linear model where the natural log of the odds ratio or logit of the categorical data is a linear combination of independent predictors $x_1, \ldots, x_n$ with intercept b_0 and weights $b_1, \ldots, b_n$, which are determined via the method of maximum likelihood as

$$\ln\left(\frac{p_e}{1-p_n}\right) = b_o + b_1 x_1 + \cdots + b_n x_n. \qquad (1)$$

To perform variable selection and the model fit, we utilize Fortran 90 code written by and provided by A. J. Miller (CSIRO 2017) that produces linear logistic

TABLE 1. Potential predictors for logistic regression algorithms to predict the formation/demise of an IR eye at lead times of 6, 12, 18, 24, and 36 h. Predictors include forecast parameters (environmental predictors) and initial conditions (IR predictors and best-track/advisory-based predictors).

Acronym	Description
	Environmental predictors (time averaged from $t = 0$ to time of the forecast)
SHRD	850- minus 200-hPa scalar wind speed difference (wind shear) calculated within a radius of 500 km
GSHR	850–200-hPa generalized wind shear calculated as the mass-weighted root-mean-square deviations of the winds from the mass-weighted deep-layer mean winds times a factor of 4 calculated in a 200–800-km annulus (Knaff et al. 2005)
SST	Sea surface temperatures at the center of the TC
OHC	Oceanic heat content between the surface and the depth of the 26°C isotherm (Shay et al. 2000 and references therein)
RH	850–500-hPa relative humidity calculated within a 200–800-km annulus
VT85	850-hPa tangential wind at 500 km from the TC center
	IR predictors
EP	Current eye probability based on LDA
C20	Percentage of IR pixels colder than 20°C within a 50–200-km annulus
C30	Percentage of IR pixels colder than 30°C within a 50–200-km annulus
C40	Percentage of IR pixels colder than 40°C within a 50–200-km annulus
C50	Percentage of IR pixels colder than 50°C within a 50–200-km annulus
IPC1	First principle component of the azimuthally averaged IR brightness temp at 0–600 km
IPC2	Second principle component of the azimuthally averaged IR brightness temp at 0–600 km
IPC3	Third principle component of the azimuthally averaged IR brightness temp at 0–600 km
IPC4	Fourth principle component of the azimuthally averaged IR brightness temp at 0–600 km
SDO	Std dev of IR brightness temp at 100–300 km
SDI	Std dev of IR brightness temp at 0–200 km
AVGO	Avg IR brightness temp at 100–300 km
AVGI	Avg IR brightness temp at 0–200 km
FR5	Deviation of IR-based TC size (R5; Knaff et al. 2014) from the climatological population as a function of TC intensity (see Knaff et al. 2016)
	Best-track/advisory-based predictors
VMAX	Current TC intensity ($t = 0$)
DV	12-h change in TC intensity
LAT	Latitude
SPD	Current storm translation speed

models by iteratively reweighted least squares. The logistic regression model, while similar to linear regression, is based on different assumptions about the relationship between dependent and independent variables. There are two primary differences: 1) since the dependent variable is binary, the conditional distribution is a Bernoulli distribution rather than a Gaussian distribution, and 2) the predicted values are probabilities of a particular outcome. Once fitted, the probability of having an eye takes the form

$$p_e = \frac{1}{[1 + e^{-(b_o + b_1 x_1 + \cdots + b_n x_n)}]}. \quad (2)$$

The measure of quality of fit is taken in terms of deviance, which is a generalization of the idea of using the sum of squares of residuals in ordinary least squares to cases where model fitting is achieved by maximum likelihood. Deviance is defined as −2 times the log-likelihood ratio of the fitted model compared to the full (i.e., perfect) model. Given the generalization, one can also define the percent deviance explained as 1 minus the ratio of the fitted model deviance to the deviance of a model containing only the intercept b_0.

Environmental conditions used in this application come from the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1999) development dataset, which is described in NOAA (2016). A number of forecast parameters were tested as potential predictors in the logistic regression scheme, which are listed and described in Table 1. Storm-centered IR imagery provides several initial conditions as potential predictors to the model. Those initial conditions are also listed and described in Table 1. Finally, best-track-based parameters were also tested in this scheme following previous research that suggested initial intensities and recent intensity trends are important to intensification. These too are listed and described in Table 1.

Several IR-based potential predictors are included in testing the scheme. Here, we use ''pixel counts'' to determine the percentage area colder than a threshold temperature as a proxy for convective vigor. The pixel

counts were calculated in a central region (50–200 km) with thresholds of −20°, −30°, −40°, and −50°C. In a similar effort to determine the spatial uniformity and overall convective vigor, brightness temperature standard deviations and averages were calculated in an inner area of 0–200 km and an outer region of 100–300 km. We are also interested in depicting the convective structure in terms of azimuthally averaged principle components. These basically represent the radial wavenumbers 0, 1, 2, and 3 as principle components 1, 2, 3, and 4. The principle components have been shown to be related to the wind field (Kossin et al. 2007; Knaff et al. 2016), the TC size (Knaff et al. 2014), and future eyewall formation (Kossin and Sitkowski 2009). Other authors (Carrasco et al. 2014; Xu and Wang 2015) have suggested that the intensification rate of cyclones is inversely related to TC size, so our final IR predictor, FR5, provides a globally homogeneous (independent of TC center) measure of TC size that accounts for variations in storm intensity. This is accomplished by calculating the IR-based TC size (R5; Knaff et al. 2014) a function of the first two IR principle components and the storm latitude and dividing that value by the intensity-based climatology of R5, as described in Knaff et al. (2016).

A forward predictor selection strategy is used for North Atlantic cases (1996–2014); adding only predictors that significantly reduce the model errors at the 99% level and to minimize the residual sum of squares while also minimizing the number of independent variables used in each model. Each potential predictor was standardized by subtracting the sample mean and dividing by the sample standard deviation, which allows for examination of both the sign and the relative contribution of each predictor. While not formally using an Akaike information criterion (Akaike 1974), this ad hoc methodology has the same aim (e.g., producing the best model with the least number of predictors). The chi-squared-based probability is also used to evaluate the model fit. For our work described in the next section, the North Atlantic EDA probabilities were used as the dependent variable for model fitting and the east Pacific EDA probabilities (1996–2014) were held out for independent testing of the resulting model. These two samples plus the 2005–14 sample from the western North Pacific (using SAB's PCN codes) are combined to create Northern Hemisphere model versions.

3. Resulting models and verification

The resulting models for forecasting eye probabilities at 6-, 12-, 18-, 24-, and 36-h lead times had between 5 and 10 predictors. These explain 80%, 66%, 52%, and 41% of the deviance at 6-, 12-, 24-, and 36-h lead times. The 18-h forecast resulted in two valid and equally well fit models (18_a and 18_b). The former explains 55% of the deviance with 5 predictors while the latter explains 60% of the deviance using 10 predictors. The key differences are the use of initial eye probabilities in the 18_a model and the use of current TC intensity (VMAX; see Table 1 for descriptions of the predictors) and several IR-based predictors in the 18_b model, so there is some independence in the calculated response. Furthermore, potential for overstating the fit or "artificial skill" in regression models increases as the number of predictors increases (Mielke et al. 1996). Table 2 provides a list of the predictors chosen at the various forecasts leads with their means and standard deviations used for the standardization. Various combinations were used, but for the sake of brevity, it is best to list all the predictors used.

TABLE 2. Predictors (independent variables) that are used in the probabilistic forecasts of IR eye probabilities, their means, standard deviations, and at which lead times they are used for model development in the North Atlantic sample.

Predictor	Mean	Std dev	Forecast lead times
SHRD	17.927	9.768	6, 12, 18_a, 18_b, 24, 36
OHC	31.579	31.115	18_b, 24, 36
VT85	7.335	3.966	18_b, 24, 36
EP	0.124	0.267	6, 12, 18_a
C50	0.376	0.268	6, 12, 18_a,18_b, 24, 36
IPC3	0.042	1.005	6, 12, 18_b
IPC4	0.004	0.985	18_b
SDO	21.688	6.645	12, 18_a, 18_b, 24, 36
FR5	1.048	0.220	18_b, 24, 36
VMAX	55.296	24.798	18_b, 24, 36
DV	2.094	9.059	6, 12, 18_a,18_b, 24, 36

A discussion of individual model combinations will follow. In general, vertical wind shear, the 500-km tangential winds at 850 hPa, and the oceanic heat content proved to be important predictors from the storm environment. These results are similar to those of environmental predictors used in the rapid intensification index (Kaplan and DeMaria 2003). Equally important is the state of the convection. Information about the current convective structure came from several predictors. It is not surprising that the current eye probability was an important predictor. The brightness temperature threshold of −50°C provided the best measure of convective vigor and the standard deviation of brightness temperatures in the 100–300-km annulus was also most predictive, suggesting rather deep and spatially uniform convection is important for future eye formation at all times. The radial wavenumbers 2 and 3 (IPC3 and IPC4), and the TC size also appear important at some leads. It is also not surprising that current intensity and recent trends are also important, as was the case in the

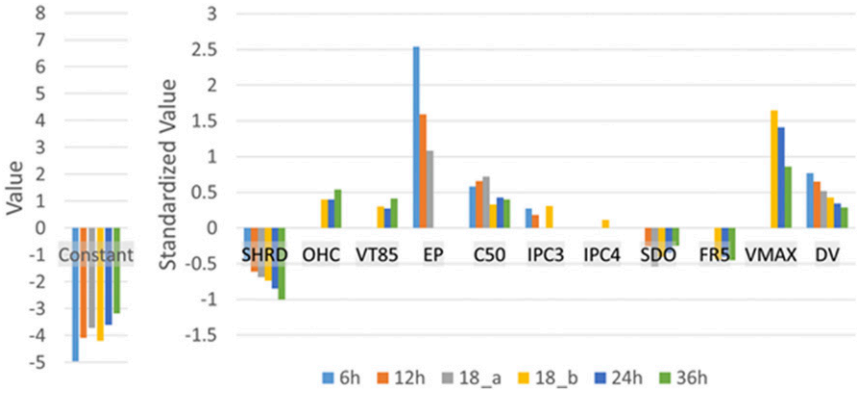

FIG. 2. Graphical depiction of the constants and standardized regression coefficients for the forecasts of eye probability at lead times of 6, 12, 18, 24, and 36 h based on the North Atlantic developmental dataset. Note that 18 h has two competitive forecast equations and both are shown here. The constants correspond to b_0 and the coefficients correspond to $b_1, b_2, \ldots, b_n$ in the logistic regression, as expressed in Eq. (2).

rapid intensification indices run in operations (e.g., Kaplan et al. 2015).

In all, six models were developed. There are two competitive and different models available at the 18-h lead; for our verification results, we will average the results of these two forecasts to produce the eye probabilities at 18 h following the concept that combining different and less dependent estimates and forecasts to form an average has long been shown to reduce uncertainty and errors (see, e.g., Student 1908; Bates and Granger 1969; Goerss 2000; Sampson et al. 2008; Sampson and Knaff 2015). Figure 2 shows the normalized regression coefficients and constants as in Eq. (2) used in each forecast model. The magnitude and sign of each predictor provides information on the relative importance of the predictor and its impact on the probability forecast. Keep in mind that in Eq. (2), the largest probabilities result when the linear combination $b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n$ equals zero, and since the constants are all negative, the linear relationship of the normalized regression coefficients is the same as their sign implies. That is, negative coefficients imply an inverse relationship with eye probabilities.

The resulting forecast equations are fairly heavily weighted to information in the IR imagery. This may be an advantage at times in anticipating rapid intensification events, as the current operational versions of the rapid intensification index (RII) are more heavily weighted to best-track information and environmental conditions. Several expected relationships also appear, including that all forecasts use some measure of current intensity and the recent intensity trends, vertical wind shear, and measures of convective vigor are statistically important at all lead times.

There are some interesting features also resulting from these statistics. The first is that for the 6- and 12-h forecasts, the persistence of eye probabilities is the most important factor. At the 18-h lead time the persistence of the current eye conditions becomes noticeably less important, and the current intensity becomes important enough to be selected as a predictor. It is also thought provoking that the initial convective vigor and spatial consistency (C50 and SDO) provide about the same amount of information at all leads, save for the 6-h forecast where SDO is not selected. The IPC3, which has been linked to future secondary eyewall formation (Kossin and Sitkowski 2009), is important for the earlier lead times, suggesting that this pattern is likely providing information about the possibility of a disruption/decrease in the eye probabilities at these leads. The effect of ocean heat content (OHC) seems to be most pronounced, influencing longer lead times, which is consistent with other statistical dynamical intensity forecasting efforts (e.g., DeMaria et al. 2005). Finally, the two predictors related to TC size, the 850-hPa tangential winds (VT85), and the IR-based TC size (FR5) also become important at and beyond the 18-h lead forecast times.

To validate the forecast schemes, we use the EDA results from the east Pacific TC basin as independent data and then compare the probabilistic forecast with two baseline forecasts. The first is climatology, which, depending on lead time, runs between 11% and 14%. The second is a simple persistence of the current eye probabilities. From this information we construct Brier skill scores (BSSs; Wilks 2006), which provide a measure of skill based on the root-mean-square errors of the forecast models versus our two baseline forecasts. We

TABLE 3. BSSs of the eye probability forecast models at various lead times based on baseline forecasts of persistence and climatology, which are based on independent forecasts of east Pacific eye formation during 1996–2014. Climatology results are 11.6%, 11.9%, 12.2%, 12.7%, and 13.9% for 6-, 12-, 18-, 24-, and 36-h forecasts, respectively. The 18-h forecasts and related statistics are also based on an equally weighted consensus of the two competitive models developed for this lead time.

Model lead time (No. of cases)	BSS (persistence)	BSS (climatology)
6 h (n = 5560)	−0.14	0.88
12 h (n = 5450)	0.35	0.80
18 h (n = 5298)	0.47	0.71
24 h (n = 5140)	0.47	0.58
36 h (n = 4563)	0.50	0.63

use the east Pacific data to demonstrate that the method works well even in a non-basin-specific and truly independent sense. In doing so, we recognize that the east Pacific basin generally has smaller storms (Knaff et al. 2014) and more favorable environmental conditions (Kaplan et al. 2010, 2015). The ultimate goal, a set of Northern Hemisphere coefficients (i.e., North Atlantic and North Pacific), is presented and discussed in section 5. The BSSs for the six forecast models are shown in Table 3. The 6-h forecast is the only time period with a negative BSS when the baseline forecast is the persistence of current conditions (i.e., worse than persistence at 6-h lead times). The rest of the forecast models show significant improvement over the two baseline forecasts with peak skill versus climatology at the 12–18-h lead times. It is noteworthy that as lead time increases, the persistence baseline forecast becomes less reliable and/or useful, so BSS versus persistence generally increases with forecast lead time.

BSS only provides a measure of relative skill versus a baseline. To provide information on the calibration of the forecasts, reliability diagrams (calibration and embedded refinement distributions) are also constructed, as shown in Fig. 3. These reliability diagrams show how the forecasts are calibrated. Generally, the shorter leads have better calibration (closer to the 1:1 line), but as we go to longer leads the forecast calibration deteriorates, and the forecasts become more overconfident (more horizontal). The forecasts also have relatively small biases (i.e., the lines are not systematically above or below the 1:1 line). The refinement distributions (note the logarithmic scale) indicate that the forecasts have high confidence (being well distributed among the potential outcomes with the largest cases at the endpoints of the distribution).

These combined results suggest that the algorithms developed to forecast IR eye probabilities using Atlantic TC cases perform well and are skillful versus the baseline forecasts of persistence and climatology. Furthermore, the forecasts appear to be well calibrated versus past validation of probabilistic intensity changes (see Kaplan et al. 2010, 2015) and show high confidence in their distribution. Finally, both the skill (i.e., positive BSSs) and reliability (relative to 1:1) are greatest at the 12–24-h forecast leads and might thus be complementary to other tools that forecast rapid TC structure and intensity changes.

4. Hurricane Matthew (2016)

We now present an example of forecasts for Hurricane Matthew (2016). To demonstrate the nature of the forecasts, we show them made every other day at 0000 UTC along with the observed eye probability for brevity (Fig. 4). All the forecasts for Matthew are available online (http://rammb.cira.colostate.edu/products/tc_realtime/archive.asp?product=eyepfcst&storm_identifier=AL142016). Hurricane Matthew demonstrated a rapid eye formation around 0600 UTC 30 September (red line), which was followed by a period of rapid intensification, intensifying from 85 to 140 kt in the next 24 h based on the working best track. The forecast made at 0000 UTC 29 September did not forecast this eye formation event, with probabilities only slightly exceeding climatology only at the 24- and 36-h lead times. This echoes the NHC forecast discussion at that time, which states the following:

> ''Matthew is forecast to move over warm waters and relatively low shear during the next few days, and these conditions should lead to gradual strengthening. There is a possibility, however, that Matthew encounters some westerly shear in the central Caribbean well south of Haiti in a day or two. This is an area climatologically unfavorable for storms to intensify, and Matthew could reduce its rate of strengthening there. After that, most of the models show a more conducive upper-level environment, and Matthew could intensify at a faster rate.''

Forecasts after that date, shown in Fig. 4, agree much better with the observations even predicting a bit of eye uncertainty that occurred late on 5 October and early on 6 October. The next to last forecast also properly forecasts the demise of the eye as Matthew weakened. Overall, such cases are encouraging. Other cases are available online (http://rammb.cira.colostate.edu/products/tc_realtime/) for 2016 TCs starting in the middle of July.

5. Northern Hemisphere coefficients

Since eye formation is a rare event, we wanted to examine the largest dataset we could assemble and develop similar logistic regression models for eye

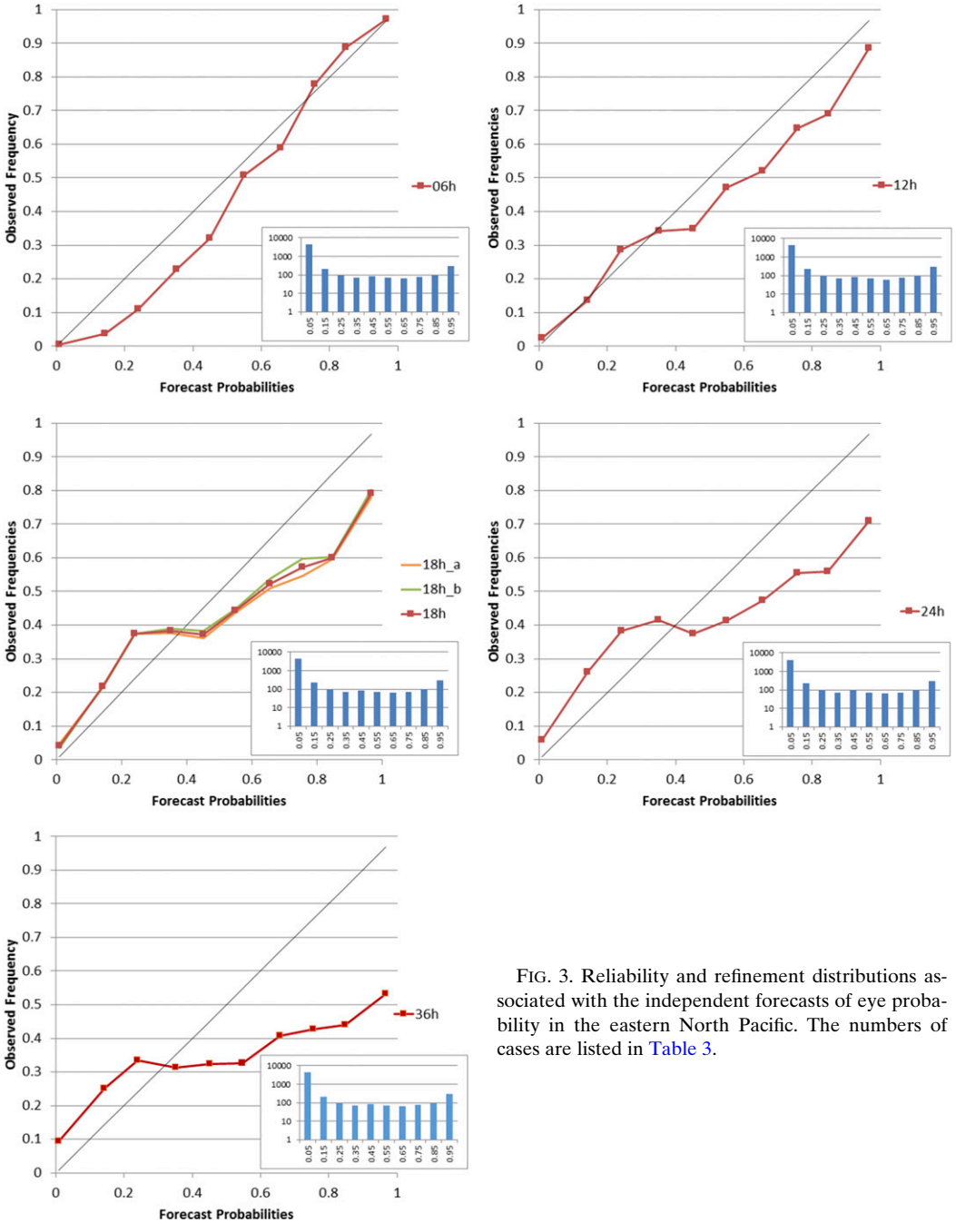

FIG. 3. Reliability and refinement distributions associated with the independent forecasts of eye probability in the eastern North Pacific. The numbers of cases are listed in Table 3.

probability. This allows us to better examine the stability of the solutions found in the North Atlantic and to develop a model that while not optimized by basin will work well across the entire Northern Hemisphere.

The dataset is compiled using SHIPS developmental data and fixes for the years 1989–2014 in the Atlantic and eastern North Pacific and for the years 2005–14 in the western North Pacific. The western North Pacific dataset was limited by the existence of hourly IR data prior to 2005. Fortunately, the sample consisted of nearly equal numbers of cases in each basin with 4673, 4962, and 4636 cases in the Atlantic, east Pacific, and west Pacific basins, respectively.

Models were constructed in a similar manner and contained the same predictors as the North Atlantic models, except for the 18_b model, where OHC and SDO were dropped and the contributions of the IPC3

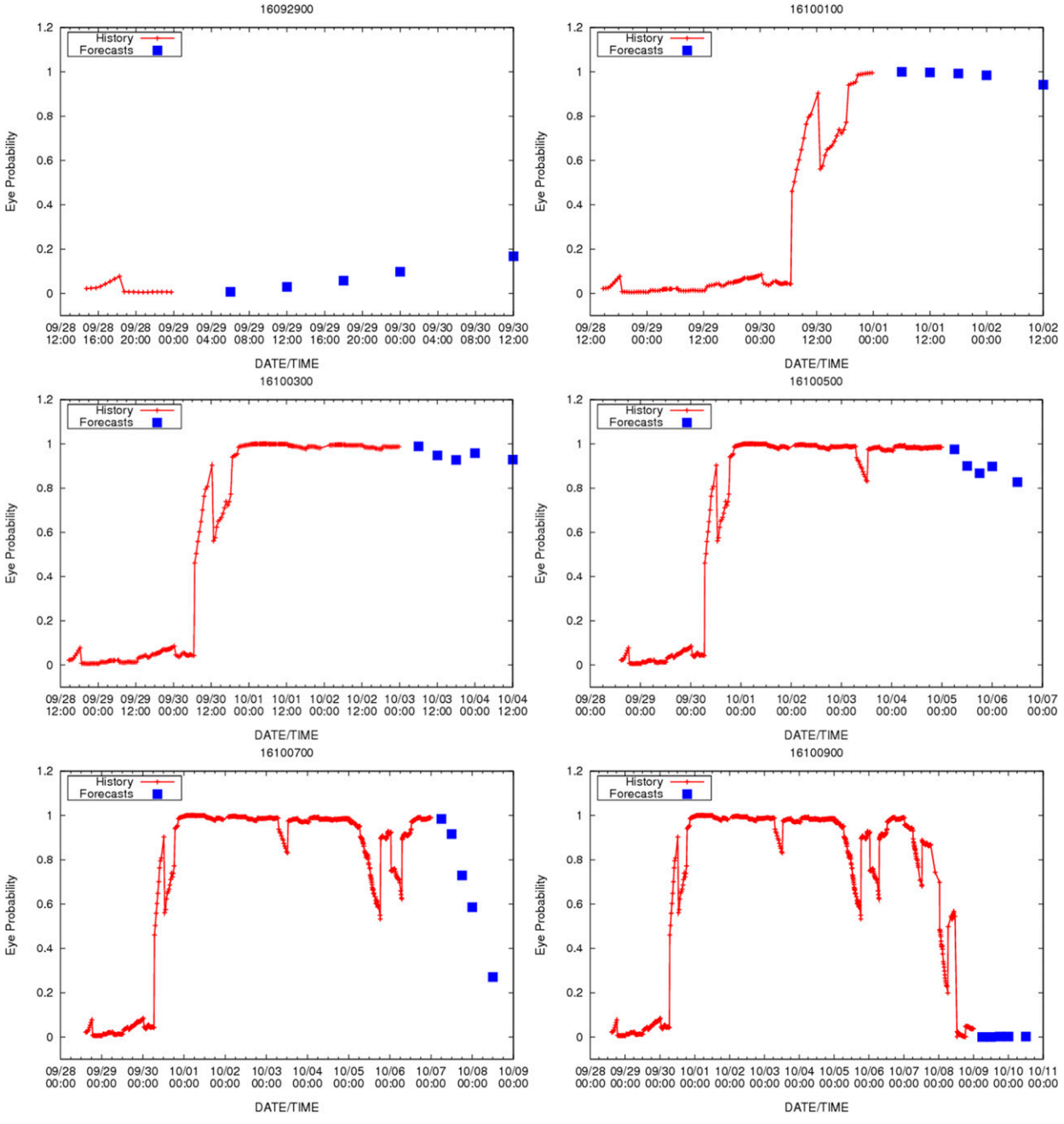

FIG. 4. The 0000 UTC forecasts made every 2 days for Hurricane Matthew (2016) starting on 29 Sep and ending on 9 Oct. The red line and points are the real-time eye probability estimates, and the blue squares are the forecasts at 6-, 12-, 18-, 24-, and 36-h lead times. Forecasts were created in real time with SHIPS large-scale diagnostic files obtained from NHC and real-time GOES-East data.

and IPC4 are very small. These dependent models explained 53%, 50%, 50%, 50%, 50%, and 43% of the deviance at 6, 12, 18_a, 18_b, 24, and 36 h, or less than in the North Atlantic. The resulting normalized weights are shown in Fig. 5. Comparing Fig. 5 with Fig. 2, one can see that the predictive influence of VWS, OHC, and C50 generally increases; the influence of EP, SDO, PC3, PC4, and VMAX decreases; and the influences of VT850 and FR5 remain very similar. Table 4 shows the statistics associated with the predictors in this much larger sample, where VWS is generally a little smaller, and OHC, VT850, and C50 are generally a little larger. The model predictor weights, however, are more similar than different when compared to the Atlantic models.

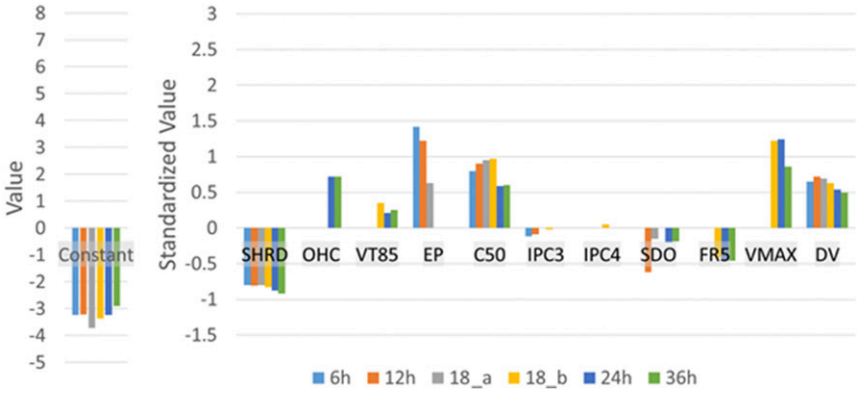

FIG. 5. As in Fig. 2, but for the Northern Hemisphere sample.

Specifically, low VWS (<~15 kt), higher OHC (>~35 kJ cm^{-2}), TCs with central cloud cover that is cold and symmetric, and smaller, more intense, and intensifying TCs are more prone to eye formation during the next 36 h.

6. Summary and discussion

This manuscript describes the development of an IR eye probability forecast scheme. The scheme was developed using an eye detection algorithm that employed a linear discriminant analysis technique to determine the probability of an eye existing in any given IR image given information about the storm center, motion, and latitude. The methodology used is logistic regression, where predictors were selected from routine information about the current storm, forecast environmental factors, and patterns and information extracted from the current IR image. Forecasts were created for 6-, 12-, 18-, 24-, and 36-h forecast leads using separate Atlantic and Northern Hemisphere developmental datasets.

The independent performance of the Atlantic basin–based model forecasts tested on independent east Pacific cases was skillful versus persistence at 12–36 h and skillful versus climatology at 6–36 h (Table 3). The calibration and reliability of the forecasts is quite good for 6–18 h, but forecasts become more overconfident (more horizontal with respect to the 1:1 line) at longer lead times (Fig. 3). The forecasts also appear to be unbiased (not systematically above or below the 1:1 line).

Northern Hemisphere models were constructed using similarly sized samples from the North Atlantic, east Pacific, and western North Pacific, and showed that having a similar set of predictors is important for anticipating IR eye formation during the next 6–36 h in the multibasin sample. While is not surprising that low-VWS, high-OHC, anomalously cold, and symmetrically deep convection near the center of TCs, as well as storms that are rather intense and have been intensifying, are more likely to form eyes in the near future, our statistics also suggest that smaller TCs are also more likely to produce eyes than larger TCs. This finding agrees with the observational results of Carrasco et al. (2014) and Xu and Wang (2015), that future intensification rates are related to TC size. These results also seem consistent with idealized modeling studies (e.g., Schubert et al. 2016), where ''the smaller vortices are better preconditioned for rapid intensification.'' Collectively, all of these results suggest that more effort in assessing the initial TC wind field and assimilating it into both statistical and dynamical models may lead to improvements in TC intensity change forecasts.

While it is understood that no operational agency predicts the formation of the TC eye, the formation of an eye represents an important structural stage where further intensification is both likely and rapid intensification more likely. In fact, the formation of an eye in IR imagery often results in at least a 15-kt increase in the

TABLE 4. Predictors (independent variables) that are used in the probabilistic forecasting of IR eye probabilities, their means, standard deviations, and at which lead times they are used for model development in the Northern Hemisphere sample.

Predictor	Mean	Std dev	Forecast lead times
SHRD	15.33	8.83	6, 12, 18_a, 18_b, 24, 36
OHC	34.98	32.53	24, 36
VT85	8.14	4.50	18_b, 24, 36
EP	0.150	0.299	6, 12, 18_a
C50	0.468	0.289	6, 12, 18_a, 18_b, 24, 36
IPC3	−0.089	1.022	6, 12, 18_b
IPC4	0.160	1.078	18_b
SDO	20.091	6.932	12, 18_a, 24, 36
FR5	1.021	0.254	18_b, 24, 36
VMAX	57.210	28.016	18_b, 24, 36
DV	2.184	10.204	6, 12, 18_a, 18_b, 24, 36

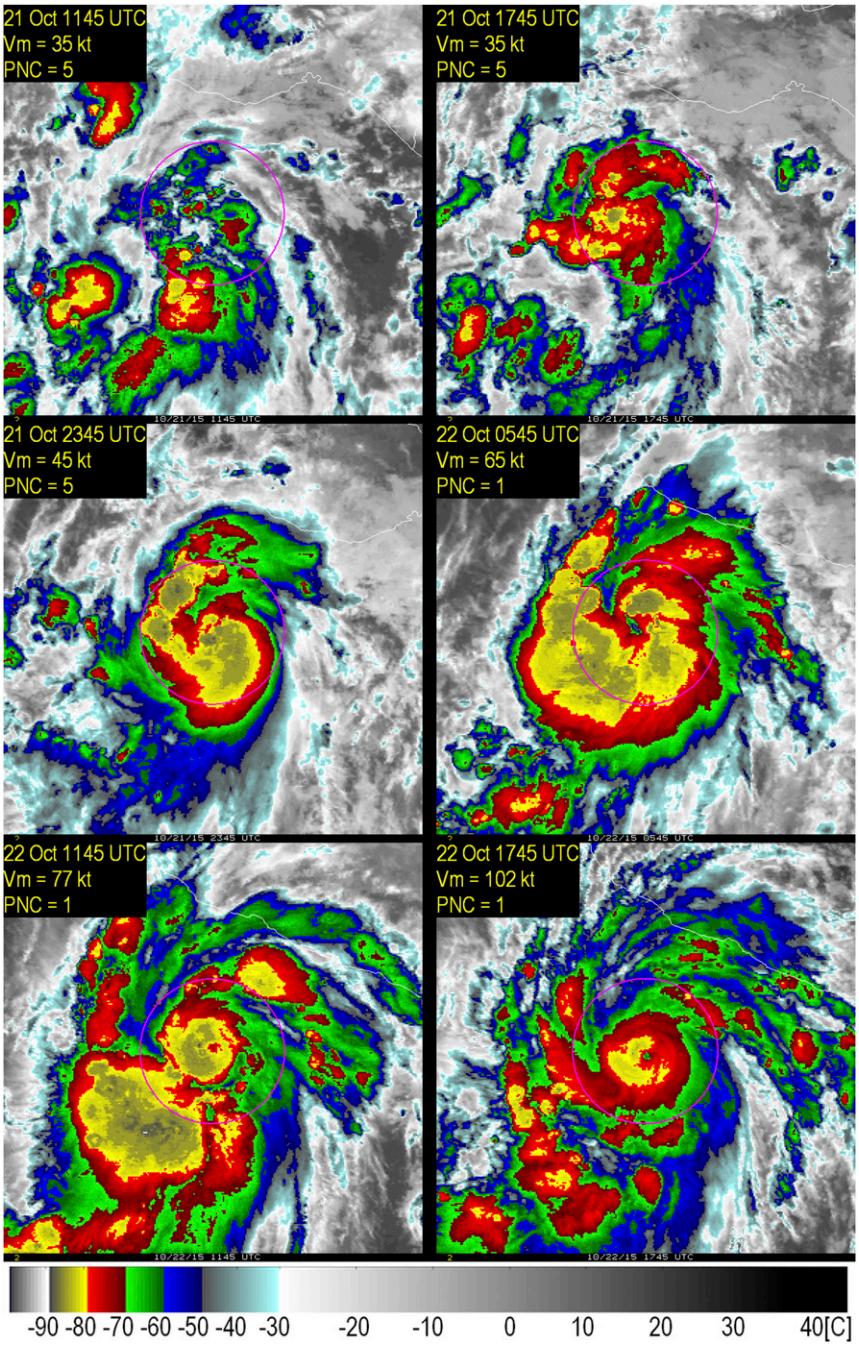

FIG. A1. A 30-h sequence of GOES-East infrared images showing the development of Hurricane Patricia (2015) in the east Pacific. The date, time, and TAFB-based intensity and PCN estimates are provided in the top left of each panel. Images are provided in a Mercator projection and a light magenta circle of 200-km radius is provided on each image to provide scale.

intensity estimate (Dvorak 1984). The disintegration of the IR eye also indicates abrupt structural changes and is typically associated with weakening. These forecasts were also created for the latter half of 2016 and were made publicly available in the hopes that they would prove useful to operational forecasters. Such forecasts may also be useful when combined with other intensity change guidance as well as observed eyewall structure

changes found in microwave imagery, where eye features typically show up when winds are around 58 kt, which is when aircraft-based radar would typically also depict an eye (Vigh et al. 2012).

Future work may include the development of eye probability forecasts using the QDA version of the EDA, the use of other statistical methods to make such forecasts, and the development of similar techniques for the Southern Hemisphere. We will also investigate if such forecasts could add to other guidance to predict rapid intensification events. Finally, we expect the next-generation Geostationary Operational Environmental Satellites will provide additional information leading to improvements of this and other techniques designed to forecast TC intensity and structure change.

Acknowledgments. This research was partially funded by both Joint Polar Satellite System (JPSS) and Geostationary Operational Environmental Satellite-R (now *GOES-16*) Risk Reduction activities under NOAA Grant NA17RJ1228 at the Cooperative Institute for Research in the Atmosphere at Colorado State University. The Hurricane Forecast Improvement Program also provided funding to maintain the SHIPS databases used for this work. We also thank Kate Musgrave and Andrea Schumacher for comments on earlier versions of this manuscript and Alan J. Miller for his software. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. government position, policy, or decision

APPENDIX

Extracting Tropical Cyclone Eye Existence Information from Routine Dvorak Fixes

Tropical cyclone position and intensity estimates (or fixes) are routinely made by several U.S. weather agencies using the Dvorak (1984) technique. For our work we used the 6-hourly fixes from SAB, which is part of the National Environmental and Satellite Data and Information Service's (NESDIS) Satellite Services Division, and TAFB, which is part of the NHC. These fixes, which also provide the IR image time and the satellite, were saved in the databases of the ATCF, whose fix formats are described in NRLMRY (2016). Part of each subjective Dvorak fix (DVTS) contains information about the position code number (PCN) and text comments. The PCN is the primary source of eye information and while the documentation suggested this information was deprecated, U.S. agencies still provide this information (B. Strahl and M. Turk 2017, personal communication). If the PCN was coded as 1 (eye/geography) or 2 (eye/ephemeris), then an eye existed in the imagery. In addition to the PCN code certain text phases in the comments were also used to determine which images have eyes: "EYE," "eye," "EMB," "emb," "BND," "bnd," "RAG," and "rag." We also partitioned eye cases into banded, ragged, and embedded features, though that information was not utilized in this study. More than 98% of the cases were classified by the PCN and then subclassified by the text information. After manually examining the imagery associated with this subjective, and possibly deprecated, information, it became obvious that the Dvorak fix was providing adequate eye classification for developing more objective techniques. Cases with intensities less than 34 kt were also excluded from the dataset. To provide the reader a bit more information, an example of six consecutive images along with their TAFB-based intensity estimates and PCNs are shown for Hurricane Patricia [2015; Kimberlain et al. (2016)] in Fig. A1. This sequence of images shows that a rapidly intensifying TC, which goes from a poorly defined system to a distinct eye in 30 h and has what appears to be a ragged eye at 0545 UTC 22 October and a pinhole eye 6 h later. Notice how the PCN corresponds to what appears to be an eye (warm spot) in the IR imagery. It is noteworthy that SAB did not classify the ragged-eye scene as having an eye at 0545 UTC 22 October, but both agencies agree on a pinhole eye at 1145 UTC 22 October.

REFERENCES

Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Autom. Control*, **19**, 716–723, doi:10.1109/TAC.1974.1100705.

Bates, J. M., and C. W. J. Granger, 1969: The combination of forecasts. *Oper. Res.*, **20**, 451–468, doi:10.1057/jors.1969.103.

Burton, A., and Coauthors, 2010: Structure and intensity change: Operational guidance. *Proc. Seventh Int. Workshop on Tropical Cyclones*, La Reunion, France, WMO/CAS/WWW, 1.5, https://www.wmo.int/pages/prog/arep/wwrp/tmr/otherfileformats/documents/1_5.pdf.

Carrasco, C., C. Landsea, and Y. Lin, 2014: The influence of tropical cyclone size on its intensification. *Wea. Forecasting*, **29**, 582–590, doi:10.1175/WAF-D-13-00092.1.

CSIRO, 2017: Software from Alan J. Miller. Commonwealth Scientific and Industrial Research Organisation, http://wp.csiro.au/alanmiller/.

DeMaria, M., 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.*, **137**, 68–82, doi:10.1175/2008MWR2513.1.

——, and J. Kaplan, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337, doi:10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvement to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, doi:10.1175/WAF862.1.

——, C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, doi:10.1175/BAMS-D-12-00240.1.

DeMaria, R. T., 2016: Automated tropical cyclone eye detection using discriminant analysis. M.S. thesis, Dept. of Computer Science, Colorado State University, 63 pp., https://dspace.library.colostate.edu/bitstream/handle/10217/170410/DeMaria_colostate_0053N_13387.pdf.

Dvorak, V. F., 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, 45 pp., http://severe.worldweather.wmo.int/TCFW/RAI_Training/Dvorak_1984.pdf.

Evans, J. L., and R. E. Hart, 2003: Objective indicators of the onset and completion of ET for Atlantic tropical cyclones. *Mon. Wea. Rev.*, **131**, 909–925, doi:10.1175/1520-0493(2003)131<0909:OIOTLC>2.0.CO;2.

Goerss, J., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193, doi:10.1175/1520-0493(2000)128<1187:TCTFUA>2.0.CO;2.

Holliday, C., and A. Thompson, 1979: Climatological characteristics of rapidly intensifying typhoons. *Mon. Wea. Rev.*, **107**, 1022–1034, doi:10.1175/1520-0493(1979)107<1022:CCORIT>2.0.CO;2.

Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin. *Wea. Forecasting*, **18**, 1093–1108, doi:10.1175/1520-0434(2003)018<1093:LCORIT>2.0.CO;2.

——, ——, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and east Pacific basins. *Wea. Forecasting*, **25**, 220–241, doi:10.1175/2009WAF2222280.1.

——, and Coauthors, 2015: Evaluating environmental impacts on tropical cyclone rapid intensification predictability utilizing statistical models. *Wea. Forecasting*, **30**, 1374–1396, doi:10.1175/WAF-D-15-0032.1.

Kimberlain, T. B., E. S. Blake, and J. P. Cangalosi, 2016: National Hurricane Center tropical cyclone report: Hurricane Patricia (EP202015). National Hurricane Center, 32 pp., http://www.nhc.noaa.gov/data/tcr/EP202015_Patricia.pdf.

Knaff, J. A., C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688–699, doi:10.1175/WAF863.1.

——, S. P. Longmore, and D. A. Molenar, 2014: An objective satellite-based tropical cyclone size climatology. *J. Climate*, **27**, 455–476, doi:10.1175/JCLI-D-13-00096.1.

——, C. J. Slocum, K. D. Musgrave, C. R. Sampson, and B. R. Strahl, 2016: Using routinely available information to estimate tropical cyclone wind structure. *Mon. Wea. Rev.*, **144**, 1233–1247, doi:10.1175/MWR-D-15-0267.1.

Kossin, J., and M. Sitkowski, 2009: An objective model for identifying secondary eyewall formation in hurricanes. *Mon. Wea. Rev.*, **137**, 876–892, doi:10.1175/2008MWR2701.1.

——, J. A. Knaff, H. I. Berger, D. C. Herndon, T. A. Cram, C. S. Velden, R. J. Murnane, and J. D. Hawkins, 2007: Estimating hurricane wind structure in the absence of aircraft reconnaissance. *Wea. Forecasting*, **22**, 89–101, doi:10.1175/WAF985.1.

Malkus, J. S., 1958: Tropical weather disturbances—Why do so few become hurricanes? *Weather*, **13**, 75–89, doi:10.1002/j.1477-8696.1958.tb02330.x.

Mielke, P. W., K. J. Berry, C. W. Landsea, and W. M. Gray, 1996: Artificial skill and validation in meteorological forecasting. *Wea. Forecasting*, **11**, 153–169, doi:10.1175/1520-0434(1996)011<0153:ASAVIM>2.0.CO;2.

Mueller, K. J., M. DeMaria, J. A. Knaff, J. P. Kossin, and T. H. Vonder Haar, 2006: Objective estimation of tropical cyclone wind structure from infrared satellite data. *Wea. Forecasting*, **21**, 990–1005, doi:10.1175/WAF955.1.

NOAA, 2016: Statistical tropical intensity forecast technique development: Developmental data. NOAA/NESDIS/Regional and Mesoscale Meteorology Branch, http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp.

Nolan, D. S., Y. Moon, and D. P. Stern, 2007: Tropical cyclone intensification from asymmetric convection: Energetics and efficiency. *J. Atmos. Sci.*, **64**, 3377–3405, doi:10.1175/JAS3988.1.

NRLMRY, 2016: ATCF fixes format, 6/10/2009. Marine Meteorology Division, Naval Research Laboratory, http://www.nrlmry.navy.mil/atcf_web/docs/database/new/newfdeck.txt.

OAR, 2017: Announcement of federal funding opportunity: FY 2017 Joint Hurricane Testbed (JHT), Hazardous Weather Testbed. NOAA/Oceanic and Atmospheric Research, 27 pp., http://www.nhc.noaa.gov/jht/NOAA-OAR-OWAQ-2017-2005004.pdf.

Olander, T. L., and C. S. Velden, 2007: The advanced Dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. *Wea. Forecasting*, **22**, 287–298, https://doi.org/10.1175/WAF975.1.

Rozoff, C. M., C. S. Velden, J. Kaplan, J. P. Kossin, and A. J. Wimmers, 2015: Improvements in the probabilistic prediction of tropical cyclone rapid intensification with passive microwave observations. *Wea. Forecasting*, **30**, 1016–1038, doi:10.1175/WAF-D-14-00109.1.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, doi:10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2.

——, and J. A. Knaff, 2015: A consensus forecast for tropical cyclone gale wind radii. *Wea. Forecasting*, **30**, 1397–1403, doi:10.1175/WAF-D-15-0009.1.

——, J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting*, **23**, 304–312, doi:10.1175/2007WAF2007028.1.

Schubert, W. H., and J. J. Hack, 1982: Inertial stability and tropical cyclone development. *J. Atmos. Sci.*, **39**, 1687–1697, doi:10.1175/1520-0469(1982)039<1687:ISATCD>2.0.CO;2.

——, C. J. Slocum, and R. K. Taft, 2016: Forced, balanced model of tropical cyclone intensification. *J. Meteor. Soc. Japan*, **94**, 119–135, doi:10.2151/jmsj.2016-007.

Shay, L., G. Goni, and P. Black, 2000: Effects of a warm oceanic feature on Hurricane Opal. *Mon. Wea. Rev.*, **128**, 1366–1383, doi:10.1175/1520-0493(2000)128<1366:EOAWOF>2.0.CO;2.

Stewart, S. R., 2016: National Hurricane Center tropical cyclone report: Hurricane Sandra (EP222015). National Hurricane Center, 18 pp., http://www.nhc.noaa.gov/data/tcr/EP222015_Sandra.pdf.

Student, 1908: The probable error of a mean. *Biometrika*, **6**, 1–25.

Velden, C. S., and Coauthors, 2006: The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. *Bull. Amer. Meteor. Soc.*, **87**, 1195–1210, https://doi.org/10.1175/BAMS-87-9-1195.

Vigh, J. L., J. A. Knaff, and W. H. Schubert, 2012: A climatology of hurricane eye formation. *Mon. Wea. Rev.*, **140**, 1405–1426, https://doi.org/10.1175/MWR-D-11-00108.1.

Weatherford, C. L., and W. M. Gray, 1988: Typhoon structure as revealed by aircraft reconnaissance. Part II: Structural variability. *Mon. Wea. Rev.*, **116**, 1044–1056, doi:10.1175/1520-0493(1988)116<1044:TSARBA>2.0.CO;2.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction*. 2nd ed. Academic Press, 627 pp.

Willoughby, H. E., 1990: Temporal changes of the primary circulation in tropical cyclones. *J. Atmos. Sci.*, **47**, 242–264, doi:10.1175/1520-0469(1990)047<0242:TCOTPC>2.0.CO;2.

Wood, K. M., and E. A. Ritchie, 2015: A definition for rapid weakening of North Atlantic and eastern North Pacific tropical cyclones. *Geophys. Res. Lett.*, **42**, 10 091–10 097, doi:10.1002/2015GL066697.

Xu, J., and Y. Wang, 2015: A statistical analysis on the dependence of tropical cyclone intensification rate on the storm intensity and size in the North Atlantic. *Wea. Forecasting*, **30**, 692–701, doi:10.1175/WAF-D-14-00141.1.